Evaluation of Zero Crossing Rate Method for Animal Ear-like Hazardous Sound Detection Device

Miho Yamada[1,*,†], Futa Goto[1,†], Takeshi Kumaki[2,†] and Kyosuke Kageyama[3,†]

[1]Graduate School of Science and Engineering Kindai University, 3-4-1 Kowakae, Higashi-Osaka, Osaka, Japan
[2]Dept. of Electronic and Computer Engineering Ritsumeikan University, 1-1-1 Noji-Higashi, Kusatsu, Shiga, Japan
[3]Dept. of Electrical, Electronic and Communication Engineering Kindai University, 3-4-1 Kowakae, Higashi-Osaka, Osaka, Japan

Abstract

Hazardous sounds exist in our daily lives. There is a possibility of encountering accidents if people cannot hear hazardous sounds. However, hearing-impaired people cannot hear these sounds and may be at risk of accidents. Therefore, some hearing-impaired people live with a hearing-assistance dog. However, the number of hearing-assistance dogs has decreased recently. Therefore, animal ear-like hazardous sound detection devices have been proposed to support hearing-impaired people. Relevant hazardous sounds include sudden sounds and approaching sounds. If these sounds can be detected and communicated to hearing-impaired people, hearing-impaired people can be aided in noticing danger around them and take action to remain safe. Furthermore, such a device can help alert hearing-impaired people about danger before it happens and to avoid it. Additionally, when the device detects hazardous sounds, it is necessary to distinguish them from surrounding noises. This paper proposes a method to distinguish the hazardous sounds using the Zero Crossing Rate (ZCR), to be implemented for recognition on the proposed device. In the experiment, a variety of sounds are recorded and categorized as sudden sounds, approaching sounds, and surrounding noises. The ZCR of each sound is calculated from the recorded sound data. The results show that the ZCR waveform can be said that it reflects the characteristics of the waveform of the sound. It is demonstrated that there are differences in the ZCR values between the sudden sounds, the approaching sounds, and the surrounding noises. Therefore, the ZCR may be a useful aid in distinguishing hazardous sounds. The ZCR can distinguish hazardous sounds regardless of their sound pressure. Also, the advantage of using ZCR is that it is independent of any specific frequency threshold when detecting. Furthermore, the ZCR is a feature that requires little computation because it is only based on counting how many times the signal changes sign. For this reason, it is suitable for real-time processing on microcomputers.

Keywords

Hazardous sound detection, Animal external ear, Hearing-impaired people, Zero Crossing Rate

1. Introduction

Hazardous sounds exist in our daily lives. There is a possibility of encountering accidents if people cannot hear hazardous sounds. However, hearing-impaired people cannot hear these sounds. In 2022, there were about 310,000 hearing-impaired people in Japan [1]. Many hearing-impaired people may be at risk of accidents if they cannot hear hazardous sounds. Therefore,

The 7th International Symposium on Advanced Technologies and Applications in the Internet of Things (ATAIT 2025), September 10-11, 2025, Kusatsu, Japan

*Corresponding author.
†These authors contributed equally.
miho.yamada97@kindai.ac.jp (M. Yamada); goto.futa1027@kindai.ac.jp (F. Goto); kumaki@fc.ritsumei.ac.jp (T. Kumaki); kageyama@ele.kindai.ac.jp (K. Kageyama)

some hearing-impaired people live with a hearing-assistance dog. The dog helps hearing-impaired people by touching them and they can detect daily hazardous sounds [2, 3]. The hearing-assistance dogs are identified by wearing an orange cape with words like "Hearing-assistance dog" [4]. Therefore, people can understand the hearing-impaired person needs support. However, the number of hearing-assistance dogs has decreased recently. Only 51 hearing-assistance dogs work in Japan currently [5]. Additionally, hearing-assistance dogs aren't widely recognized in society and they are refused entry to public facilities such as restaurants [6]. Therefore, animal ear-like hazardous sound detection devices have been proposed to support hearing-impaired people [7, 8, 9, 10, 11, 12]. Such devices detect hazardous sounds, and notify the hearing-impaired people of danger sources.

Relevant hazardous sounds include sudden sounds and approaching sounds. Example of sudden sounds include alarms, bicycle bells, something falling down, and so on. These sounds are signs of danger and environmental change. On the other hand, approaching sounds are sounds made by approaching trains, bicycles, cars, and so on. These sounds also indicate potential danger. They both help us notice danger around us quickly. So, it is important to pay attention to these sounds and take action to remain safe in daily life. Surrounding noises (aka, ambient noise) are the background sounds that you hear in everyday places. This noise is always there, even if you don't notice it. Surrounding noises can make it difficult to hear the sounds you need to be aware of. If these various sounds can be detected and communicated to hearing-impaired people, hearing-impaired people can be aided in noticing danger around them and take action to remain safe. Furthermore, it can help hearing-impaired people know about danger before it happens and avoid it. Also, when the device detects hazardous sounds, it is necessary to distinguish them from surrounding noises. This paper proposes a method to distinguish the hazardous sounds to be implemented for recognition on the proposed device. This study specifically investigates a distinguishing approach based on the Zero Crossing Rate (ZCR).

The rest of the paper is organized as follows. Section 2 explains existing technologies. Section 3 describes the animal ear-like hazardous sound detection device. Section 4 outlines the Zero Crossing Rate (ZCR). Section 5 explains experimental method consists of the prototype of animal ear-like hazardous sound detection device, recording environment, and calculation of the ZCR. Section 6 shows the ZCR values as experimental results, and Section 7 gives our conclusions.

2. Existing technologies

There are four existing technologies that are related to either detecting hazardous sounds or supporting hearing-impaired people. The first system identifies hazardous sounds and situations through clustering based on the complete linkage method and probabilistic modeling of daily environmental sounds [13]. This system requires a lot of calculations to process large volumes of data. The second system monitors the elderly through sound. It supports health management and danger detection by analyzing indoor acoustic environments using microphone sensors. In this system, acoustic events are extracted from audio signals recorded by microphones through machine learning techniques, providing insights into daily living patterns [14]. However, the use of machine learning can lead to increased processing demands and higher power consumption. As a result, these systems are difficult to implement on a mobile device. The third system distinguishes between daily sounds and non-daily sounds using a microphone array [15]. This system is fixed in place and installed at a height of about 3 meters, making it unsuitable for portable use. The fourth system is a hearing aid. Hearing aids support hearing-impaired people by amplifying sounds. However, they have certain drawbacks, such as high cost, and users may experience discomfort due to muffled sound quality or the echo of their own voice.

3. Animal ear-like hazardous sound detection device

The animal ear-like hazardous sound detection device is explained in this section. An overview of this device is shown in Fig. 1. Hazardous sounds assist people to notice danger. The processing flow of the proposed device is shown below.

1. The device continuously gets sound from the microphone sensor attached an animal external attachment (Fig. 1 (i)).
2. The acquired sound is analyzed with a built-in microcomputer and is judged to be a hazardous sound or not (Fig. 1 (ii)).
3. If the sound is judged to be a hazardous sound, the device notifies the hearing-impaired people using methods such as LEDs and vibrations (Fig. 1 (iii)).

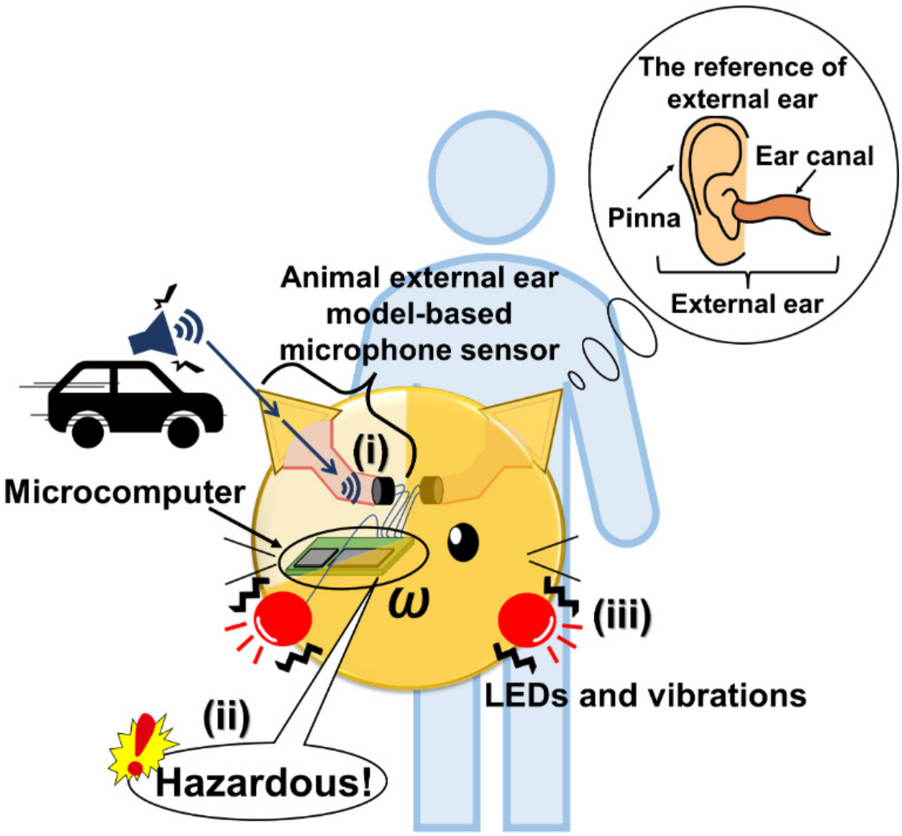

Figure 1: Overview of the animal ear-like hazardous sound detection device

The animal external attachment is shaped somewhat like the external ear of an animal. Because animals have two ears, they are able to judge the direction and source of sounds. Animals can detect enemies and danger in real time by using their ears. Their hearing abilities vary widely across species. That said, animals have evolved their ears for sound localization, which helps them detect and respond quickly to enemies and danger [16]. The external ear is made up of two structures: the pinna and the ear canal. The pinna has a sound collection effect and the ear canal has a resonance effect [17]. It is believed that utilizing the shape of the external ear in an external attachment could improve the detection of hazardous sounds. The sound collection effect of the pinna is examined. The experiment compares conditions with and without an external attachment that models an animal's external ear. In addition, the differences in sound pressure received are evaluated depending on the shape of the external attachment. Also, the variation in the observed frequency characteristics due to the length of the ear canal in the external attachment are also examined.

4. Zero Crossing Rate (ZCR)

The zero crossing is the number of times the sound waveform crosses the zero level. Specifically, the Zero Crossing Rate (ZCR) is defined as the proportion of zero crossings within a single frame like Fig. 2. The ZCR is used to distinguish voiced and unvoiced signals [18]. When the ZCR is high, the sound contains more noise. The ZCR (m) of frame m is defined as equation. (1).

$$\mathrm{ZCR}(m)[\%] = \frac{100}{N} \sum_{n=1}^{N-1} \frac{1}{2} \left| \mathrm{sgn}(x[n+mH]) - \mathrm{sgn}(x[n-1+mH]) \right| \tag{1}$$

The frame length is N samples, and the frame shift length is H samples. The frame shift is the number of samples between frame starts. The function $\mathrm{sgn}(x)$ in equation. (1) is defined equation. (2).

$$\mathrm{sgn}(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases} \tag{2}$$

In equation (1), the following term appears:

$$\frac{1}{2} \left| \mathrm{sgn}(x[n+mH]) - \mathrm{sgn}(x[n-1+mH]) \right|$$

This term takes the value 1 if the signs of two consecutive samples differ, and 0 otherwise. The factor $\frac{1}{2}$ is used to make the term equal 1 when a zero crossing occurs, because the difference of the sign values becomes $\pm\, 2$ in that case. Adding this term from $n = 1, \ldots, N-1$ gives the total number of zero crossings in the frame. Dividing this number by N normalizes it according to the frame length. The result of equation. (1) is multiplied by 100 to express the ZCR as a percentage.

Consequently, the ZCR provides a simple and effective measure of the spectral characteristics of the signal. It is widely used in speech processing tasks such as voiced and unvoiced classification and noise estimation. In addition, the ZCR is considered to be effective for distinguishing

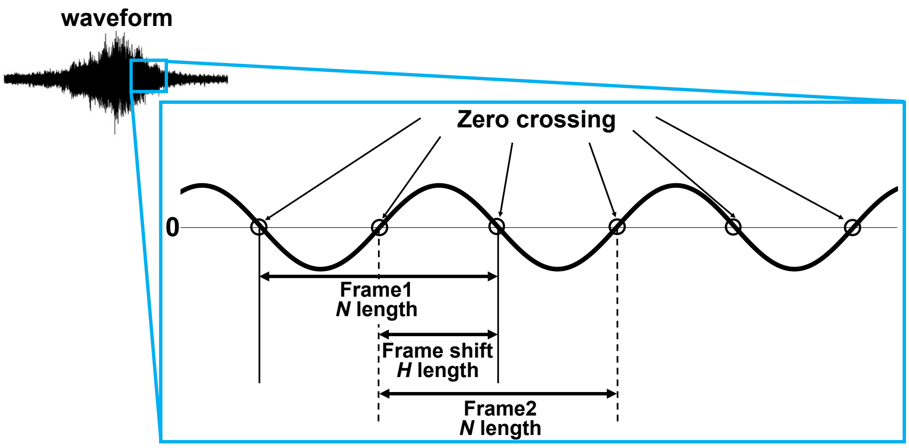

Figure 2: Zero Crossing Rate algorithm

such as the sudden sounds, the approaching sounds, and the surrounding noises. Its applicability to these sounds is examined in this study.

5. Experimental method

5.1. The prototype of animal ear-like hazardous sound detection device

A prototype of the animal ear-like hazardous sound detection device is used in the experiment. An overview of the prototype device's external appearance and inner structure is shown in Fig. 3 (a). The prototype device includes a SPRESENSE microcomputer, consisting of a main board (CXD5602PWBMAIN1) and an extension board (CXD5602PWBEXT1) [19], to which a microphone sensor (KY-037) and an LED (LEDM226-12B5) are connected. Furthermore, an external attachment modeled after a cat ear-shaped external ear is mounted on the microphone sensor. Cat's ears have been selected as a model because they are known for their excellent hearing. Generally, a cat can hear sounds from approximately 0.05 $\sim$ 65 kHz [20]. The external attachment inspired by the shape of a cat's ear is created using a 3D printer. TPU (thermoplastic polyurethane) is used to reproduce the softness of a real cat's ear. The prototype device is powered by a mobile battery, enabling continuous operation for the duration of the battery life. Also, the prototype device is lightweight and compact. Its charming design is intended to accept by the public.

5.2. Recording environment

The sudden sounds, the approaching sounds, and the surrounding noises are recorded using the prototype device. The recording environment is outdoors and included ambient and surrounding environmental noise, reflecting the actual environment. (Fig. 3 (b)). The prototype device is attached to a backpack. The recorder is wearing the backpack while recording the sounds. Several types of the sudden sounds, the approaching sounds, and the surrounding noises are

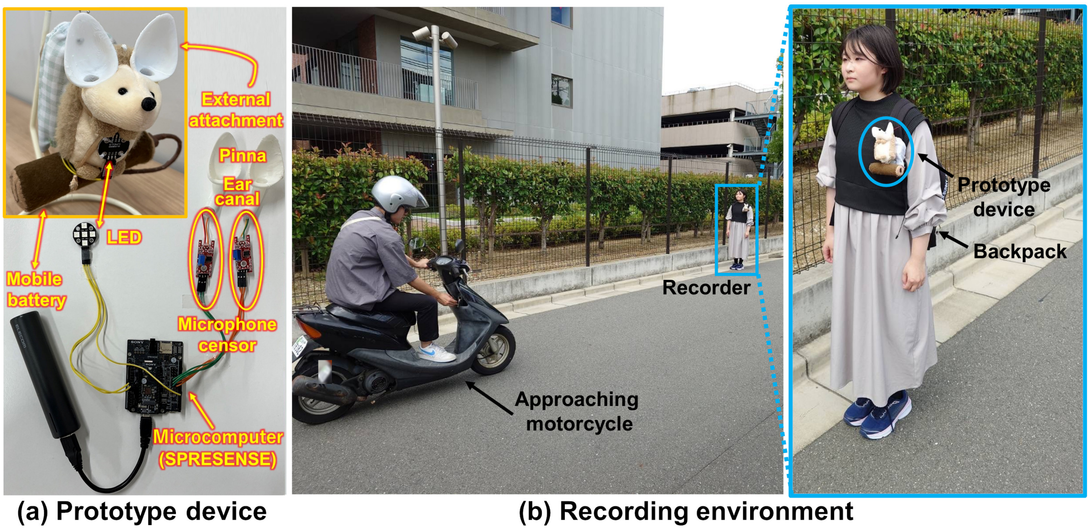

Figure 3: The prototype device and recording environment

recorded. The recorded sounds include the following. The sudden sounds are motorcycling horn sounds and bicycle bell sounds. The approaching sounds are approaching motorcycle sounds and approaching car sounds. These sounds are selected because they are perceived as dangerous in daily life and, within each category, exhibit features common to other hazardous sounds. The surrounding noises are defined as the sound recorded outdoors when no specific sounds are present. In total, there are 10 samples for each type of sound. Also, the recording is conducted at a sampling frequency of 48 kHz.

5.3. Calculation of the ZCR

The ZCR is calculated from the recorded data of the sudden sounds, the approaching sounds, and the surrounding noises. The calculation is applied to particular sound sections within the recorded data of each sound. The ZCR of each sound is calculated using equation (1) for each frame. In this study, equation. (1) is calculated using a frame length N=1,024 and a frame shift length H=512, resulting in a 50% overlap between frames. A frame length of 1,024 is chosen to balance temporal resolution and stability, and a 50% overlap is used to smoothly track sudden changes. Then, the average of the ZCR is calculated for each sound section. Finally, the ZCR values are compared between the sudden sounds, the approaching sounds, and the surrounding noises.

6. Experimental result

Figure 4 shows the result of the waveform and the ZCR waveform in the sudden sounds, the approaching sounds, and the surrounding noises. A motorcycling horn sound is used to show the sudden sounds. An approaching motorcycle sound is used to show the approaching sounds. Ambient noise found outside is used to show the surrounding noises. In Fig. 4 (i) (b), the ZCR waveform of the sudden sounds shows a sharp increase at the onset of the motorcycling horn

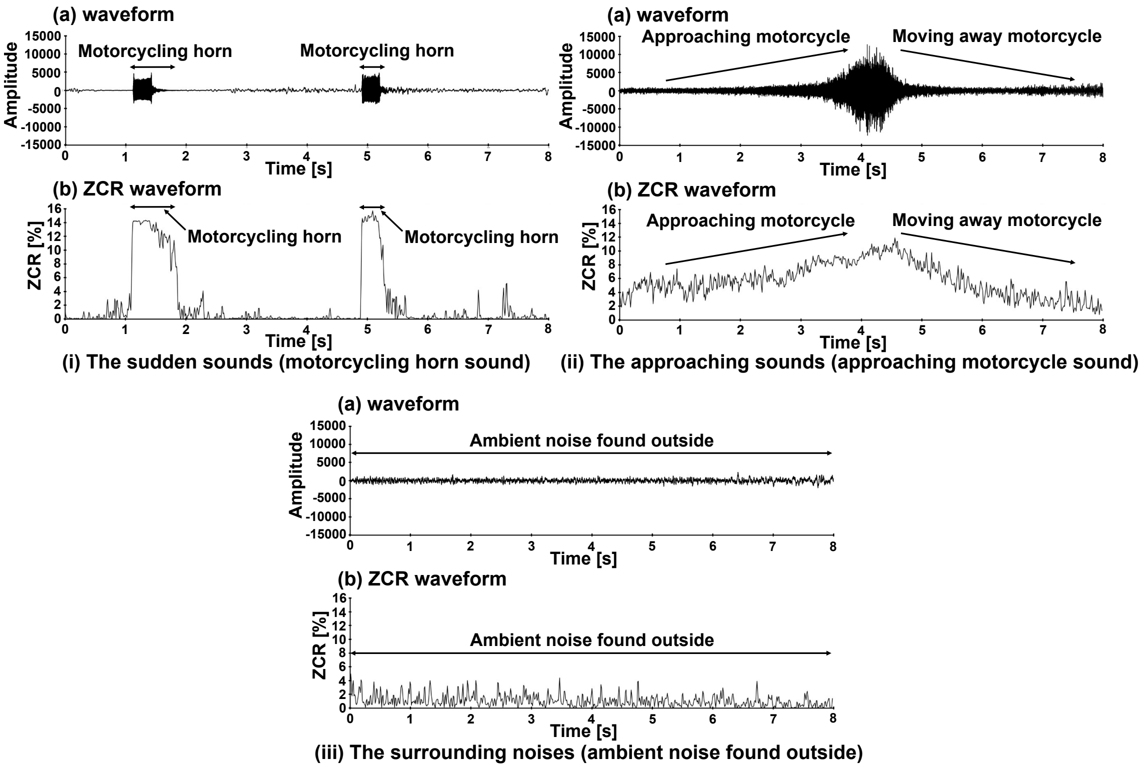

Figure 4: The waveform and the ZCR waveform of each type of sound

sound and a rapid decrease when the motorcycling horn sound ended. In Fig. 4 (ii) (b), the ZCR waveform of the approaching sounds shows a gradual increase as the motorcycle approaches and a gradual decrease as it moves away. This can be considered a result of the Doppler effect. In Fig. 4 (iii) (b), the ZCR waveform of the surrounding noises shows small fluctuations, but an increasing or decreasing trend is not observed. These results show that the ZCR waveform reflects the characteristics of the waveform of the sound itself. Also, Fig. 4 (i) and (ii) show that an increase of sound amplitude does not lead to an increase of the ZCR.

Table. 1 shows the average of the ZCR from particular sound sections of 10 samples for each type of sound, including sudden sounds, approaching sounds, and surrounding noises. The average of the ZCR for the sudden sounds is 14.40 %. The average of the ZCR for the approaching sounds is 6.70 %. The average of the ZCR for the surrounding noises is 1.48 %. These results show that there are differences in the ZCR values between the sudden sounds, the approaching sounds, and the surrounding noises. Furthermore, it is confirmed that the ZCR can be obtained in noisy environments without being affected by noise. Therefore, the ZCR may be a useful aid in distinguishing hazardous sounds. The ZCR can distinguish hazardous sounds regardless of their sound pressure. Also, hazardous sounds vary greatly, making it difficult to limit detection to specific frequencies. However, the ZCR has the advantage of being independent of any specific frequency threshold when detecting. Furthermore, the method to

Table 1
The average of the ZCR for each type of sound

Sound source	The average of the ZCR of each sound [%]
The sudden sounds	14.40
The approaching sounds	6.70
The surrounding noises	1.48

distinguish the hazardous sounds using the ZCR is considered applicable without the need for additional filtering processes. The ZCR is a feature that requires little computation because it is only based on counting how many times the signal changes sign. For this reason, it is suitable for real-time processing on microcomputers.

7. Conclusion

This paper proposes a method to distinguish the hazardous sounds using the Zero Crossing Rate (ZCR), to be implemented for recognition on the proposed device. For the experiment, three types of sounds were investigated:1) sudden sounds; 2) approaching sounds; and 3) surrounding noises are recorded using the prototype of an animal ear-like hazardous sound detection device. Subsequently, the ZCR of each sound is calculated from the recorded sound data. The results show that the ZCR waveform reflects the characteristics of the waveform of the sound itself. Furthermore, there are differences in the ZCR values between the sudden sounds, the approaching sounds, and the surrounding noises. Therefore, the ZCR may be a useful aid in distinguishing hazardous sounds. The ZCR can distinguish hazardous sounds regardless of their sound pressure. Also, the advantage of using ZCR is that it is independent of any specific frequency threshold when detecting. Furthermore, the ZCR is a feature that requires little computation because it is only based on counting how many times the signal changes sign. For this reason, it is suitable for real-time processing on microcomputers.

In the future, sounds that are not used in the experiment will also be tested. For example, emergency sirens and falling objects. In addition, recording is conducted in environments with rain and wind, as well as in situations where many sounds overlap. It is verified whether detection correctly under these conditions. Furthermore, this method is planned to be implemented in animal ear-like hazardous sound detection devices. Finally, the goal is to develop a device that can more accurately detect hazardous sound and more broadly assist the hearing-impaired people in real-world conditions.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] https://www.mhlw.go.jp/toukei/list/dl/seikatsu_chousa_c_r04.pdf (Accessed, Jun. 2025)

[2] https://www.hojyoken.or.jp/what/hearingdog (Accessed, Jun. 2025)
[3] Miyuki Mizukoshi, “Hearing dogs - an aid for information communication,” Journal of the Acoustical Society of Japan, Vol. 72, No. 6, pp. 332-339, 2016.
[4] https://www.hearingdogjp.org/about/workandrole (Accessed, Jun. 2025)
[5] https://www.jsdrc.jp/hojoken/chodoken_suu/ (Accessed, Jun. 2025)
[6] https://grapee.jp/99344 (Accessed, Jun. 2025)
[7] Miho Yamada, Takeshi Kumaki, and Kyosuke Kageyama, “Proposal of hazard sound detection system based on the external ear model,” IPSJ SIG Technical Report, Vol. 2023-EMB-64, No. 9, pp. 1–4, Nov., 2023.
[8] Miho Yamada, Takeshi Kumaki, and Kyosuke Kageyama, “Proposal of External-ear Model-based Mobile Device for Hazard-sound Detection,” The 39th International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC2024), Jul., 2024.
[9] Miho Yamada, Takeshi Kumaki, and Kyosuke Kageyama, “Verification of Animal External Ear Model-based Microphone Sensor for Hazardous Sound Detection,” The 6th International Symposium on Advanced Technologies and Applications in the Internet of Things (ATAIT2024), Aug., 2024.
[10] Miho Yamada, Takuto Torigai, Futa Goto, Takeshi Kumaki, and Kyosuke Kageyama, “Validation of Ear Canal Resonance for Hazardous Sound Detection Method using Animal External Ear,” The 2024 IEICE Society Conference, A-8-05, Sep., 2024.
[11] Miho Yamada, Futa Goto, Takeshi Kumaki, and Kyosuke Kageyama, “Development of Animal External Ear Model-used Prototype Device for Surrounding Hazardous Sounds Detection,” IPSJ SIG Technical Report, Vol. 2024-EMB-67, No. 6, pp. 1–2, Nov., 2024.
[12] https://www.khc.or.jp/ology/tecno2024.html (Accessed, Jun. 2025)
[13] K. Suzuki, M. Tanaka, and H. Inoue, “Clustering and modeling of hazardous or non-hazardous sound in daily life,” 2012 Proceedings of SICE Annual Conference (SICE), Japan, pp. 1991–1995, 2012.
[14] https://www.fujitsu.com/jp/about/resources/case-studies/trends/pdf/cs-201710-connectedlife.pdf. (Accessed, Jun. 2024)
[15] M. Kawamoto, F. Asano, and K. Kurumatani, “A Security Monitoring System of Detecting Unusual Sounds and Hazardous Situations by Sound Environment Measurement Using Microphone Arrays,” IPSJ SIG Technical Report, Vol. 2008-UBI-19, No. 66, pp. 19–26, Jul. 2008.
[16] Ryo Oda, “Evolutionary biology of binaural processing,” Journal of the Acoustical Society of Japan, Vol. 58, No. 3, pp. 193-198, 2002.
[17] J. Melloui, O. Bouattane, and J. Bakkoury, “Study of the effect of a cause of tinnitus on the resonant frequency of the outer ear,” 2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), pp. 1–4, 2020.
[18] N. Chauhan, T. Isshiki and D. Li, "Speaker Recognition Using LPC, MFCC, ZCR Features with ANN and SVM Classifier for Large Input Database," 2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS), Singapore, 2019.
[19] https://www.sony-semicon.com/ja/products/spresense/index.html (Accessed, Jun. 2025)
[20] https://www.signia.net/ja-jp/blog/local/ja-jp/cat-dog-human-hearing/ (Accessed, Jun. 2025)